



SEQUENCE LISTING

<110> INGRAM, LONNIE
ZHOU, SHENGDE

<120> METHODS AND COMPOSITIONS FOR SIMULTANEOUS SACCHARIFICATION AND
FERMENTATION

<130> 49950-59776US

<140> 09/885,297

<141> 2001-06-19

<150> PCT/US01/19690

<151> 2001-06-19

<150> 60/214,137

<151> 2000-06-26

<150> 60/219,913

<151> 2000-07-21

<160> 24

<170> PatentIn version 3.2

<210> 1

<211> 450

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
promoter sequence

<400> 1

ctttttcggc atgagcaacc aacattttca aggtatcatc ctgatgcgca atatcggcat	60
cgggttagcca taaccatttt acctgtccgg cggccttaat accttgatca gatgggttcgt	120
ggtgttggtta ccttgccgaa gggcacccgt aaaaatgttc gcgtcgggtgt tttcgcccg	180
ggcccgaaag ctgaagaagc taaagctgct ggtgcagaag ttgtcggcgc agaagacctg	240
atggaagcca ttcagggcgg cagcattgat ttcgatcgtg atgcccttta tactgaaatt	300
gccttgcgct gccataatga agcagcctcc ggtgttttgg cagatttaag cgctgcctga	360
ttttcgtgat cctctagagt ctatgaaatg gagattcatt tatgcctctc tcttattcgg	420
ataaccatcc agtcacccgc aagcttggcc	450

<210> 2

<211> 1509

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
expression vector

<400> 2

gatcaaccgg caatttatcc acggcatcaa attcgatctg tcttttcccg tatcattggc	60
aataccggca ttctgattac aggccgtggt ttgaatgcgg tatgcagttt tgtctatgtc	120
gcatggacat cccagacatt gggattgaac ctgtttggtg tcatgctttt gattacgact	180
tttgctaccc tgatttcgga tattaccgtg tttcagtcac ggcaaaccct gctgcattac	240
ggttcaaaag cttttcagga aaaagatttt aaccaatttg atgatgtcct tgcccttttg	300
atcagagccg attttttttag tgcggcgata ggtatgttgg tagggttagg cggatatctg	360
attttaggca cttcaagatt gggatggcct gccgaggtca agccagatgc cttgctttgt	420
atgctgatta tactttttat gaatatcggc tggccaacc gggatgttgc ggctgtgtaa	480
ccgctttaa ctggtcacta tttatgagtt tattacgacc tgcgtcagaa ccggagggtg	540
tggcattggg tattggcttc atatgccttt ggggtatttt ttgtttatat ggtgcctgac	600
gcaattcacg ctttttgtca cctgtagtta cgctggcatt tatctctttc accaatatac	660
ggagcgagca tttccgataa gaaaaatatt tcagagaaaa acgcccgttg aagggatgtg	720
gaaattcact ttaagcgtca gttttaatga aatcctagac tccattttcc agcaggggtg	780
cacccttgct attggtagct cactgggggc tggggaagcc gctgtctatc gggtcgcgcg	840
ccagattagt aacggtttat ccaaaccagc acagatgatg atcggctaac atgcatccac	900
cggcagcacc ggccgtttta tgcttgggat tattgatatg ccgaaaagga tacaacatct	960
ggaagaaaaa gacgaaggcc ggaataagcg ccattctgc aaaattgtta caacttagtc	1020
gcgccatcag ggaatgaaaa atcaatccgt ctttttcggc atgagcaacc aacattttca	1080
aggtatcatc ctgatgcgca atatcggcat cggttagcca taaccatttt acctgtccgg	1140
cggccttaat accttgatca gatgggtcgt ggtgttggtta ccttgccgaa gggcaccggg	1200
aaaaatgttc gcgtcgggtg tttcgccgtg ggcccgaaag ctgaagaagc taaagctgct	1260
ggtgcagaag ttgtcggcgc agaagacctg atggaagcca ttcagggcgg cagcattgat	1320
ttcgatcgtg atgcccttta tactgaaatt gccttgcgct gccataatga agcagcctcc	1380
ggtgttttgg cagattttaag cgctgcctga ttttcgtgat cctctagagt ctatgaaatg	1440
gagattcatt tatgcctctc tcttattcgg ataaccatcc agtcatccgc aagcttggcc	1500
gtaatccat	1509

<210> 3
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic primer

 <400> 3
 cgaattcctg ccgaagttaa ttagcca 27

<210> 4
 <211> 31
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic primer

 <400> 4
 aaggatcctt ccaccagcta tttgtagtg a 31

<210> 5
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic primer

 <400> 5
 agaattctgc cagttggttg acgatag 27

<210> 6
 <211> 30
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic primer

 <400> 6
 caggatcccc tcaagtcact agttaaactg 30

<210> 7
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic primer

 <400> 7
 taatacgact cactataggg 20

<210> 8
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 8
cacgacgttg taaaacgac 19

<210> 9
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 9
taacaatttc acacagga 18

<210> 10
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 10
gactggatgg ttatccgaat aagagagagg 30

<210> 11
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
nucleotide linker

<400> 11
ggcgcgcc 8

<210> 12
<211> 11544
<212> DNA
<213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic nucleotide
 vector sequence

<220>
 <221> CDS
 <222> (1452)..(2735)

<220>
 <221> modified_base
 <222> (3282)..(4281)
 <223> a, c, g, t, unknown, or other

<220>
 <221> CDS
 <222> (4916)..(5776)

<220>
 <221> CDS
 <222> (7061)..(8251)

<220>
 <221> modified_base
 <222> (9476)..(11544)
 <223> a, c, g, t, unknown, or other

<400> 12
 gatcaaccgg caatttatcc acggcatcaa attcgatctg tcttttcccg tatcattggc 60
 aataccggca ttctgattac aggcctgtgt ttgaatgcgg tatgcagttt tgtctatgtc 120
 gcatggacat cccagacatt gggattgaac ctgtttgggtg tcatgctttt gattacgact 180
 tttgctaccc tgatttcgga tattaccctgt tttcagtcac ggcaaacctt gctgcattac 240
 ggttcaaaag cttttcagga aaaagatttt aaccaatttg atgatgtcct tgccttttgc 300
 atcagagccg atttttttag tgcggcgata ggtatgttgg tagggttagg cggatatctg 360
 attttaggca cttcaagatt gggatggcct gccgaggtca agccagatgc cttgctttgt 420
 atgctgatta tactttttat gaatatcggc tgggtccaacc gggatgttgc ggctgtgtaa 480
 ccgcttttaa ctggtcacta tttatgagtt tattacgacc tgcgtcagaa ccggagggtg 540
 tggcattggg tattggcttc atatgccttt ggggtatttt ttgtttatat ggtgcctgac 600
 gcaattcacg ctttttgtca cctgtagtta cgctggcatt tatctctttc accaatatac 660
 ggagcgagca tttccgataa gaaaaatatt tcagagaaaa acgcccgttg aagggatgtg 720
 gaaattcact ttaagcgtca gttttaatga aatcctagac tccattttcc agcaggggtg 780
 cacccttgct attggtagct cactgggggc tggggaagcc gctgtctatc gggtcgcgcg 840
 ccagattagt aacgggtttat ccaaaccagc acagatgatg atcggctaac atgcatccac 900
 cggcagcacc ggccgtttta tgcttgggat tattgatatg ccgaaaagga tacaacatct 960

ggaagaaaaa gacgaaggcc ggaataagcg cccattctgc aaaattgtta caacttagtc 1020
 gcgccatcag ggaatgaaaa atcaatccgt ctttttcggc atgagcaacc aacattttca 1080
 aggtatcatc ctgatgcgca atatcggcac cggttagcca taaccatttt acctgtccgg 1140
 cggccttaat accttgatca gatgggttcgt ggtgttggtta ccttgccgaa gggcaccggc 1200
 aaaaatgttc gcgtcgggtgt ttctgcccgt ggcccgaaag ctgaagaagc taaagctgct 1260
 ggtgcagaag ttgtcggcgc agaagacctg atggaagcca ttcagggcgg cagcattgat 1320
 ttcgatcgtg atgcccttta tactgaaatt gccttgcgct gccataatga agcagcctcc 1380
 ggtgttttgg cagatttaag cgctgcctga ttttcgtgat cctctagagt ctatgaaatg 1440
 gagattcatt t atg cct ctc tct tat tgc gat aac cat cca gtc atc gat 1490
 Met Pro Leu Ser Tyr Ser Asp Asn His Pro Val Ile Asp
 1 5 10
 agc caa aaa cac gcc cca cgt aaa aaa ctg ttt cta tct tgt gcc tgt 1538
 Ser Gln Lys His Ala Pro Arg Lys Lys Leu Phe Leu Ser Cys Ala Cys
 15 20 25
 tta gga tta agc ctt gcc tgc ctt tcc agt aat gcc tgg gcg agt gtt 1586
 Leu Gly Leu Ser Leu Ala Cys Leu Ser Ser Asn Ala Trp Ala Ser Val
 30 35 40 45
 gag ccg tta tcc gtt agc ggc aat aaa atc tac gca ggt gaa aaa gcc 1634
 Glu Pro Leu Ser Val Ser Gly Asn Lys Ile Tyr Ala Gly Glu Lys Ala
 50 55 60
 aaa agt ttt gcc ggc aac agc tta ttc tgg agt aat aat ggt tgg ggt 1682
 Lys Ser Phe Ala Gly Asn Ser Leu Phe Trp Ser Asn Asn Gly Trp Gly
 65 70 75
 ggg gaa aaa ttc tac aca gcc gat acc gtt gcg tgc ctg aaa aaa gac 1730
 Gly Glu Lys Phe Tyr Thr Ala Asp Thr Val Ala Ser Leu Lys Lys Asp
 80 85 90
 tgg aaa tcc agc att gtt cgc gcc gct atg ggc gtt cag gaa agc ggt 1778
 Trp Lys Ser Ser Ile Val Arg Ala Ala Met Gly Val Gln Glu Ser Gly
 95 100 105
 ggt tat ctg cag gac ccg gct ggc aac aag gcc aaa gtt gaa aga gtg 1826
 Gly Tyr Leu Gln Asp Pro Ala Gly Asn Lys Ala Lys Val Glu Arg Val
 110 115 120 125
 gtg gat gcc gca atc gcc aac gat atg tat gtg att att gac tgg cac 1874
 Val Asp Ala Ala Ile Ala Asn Asp Met Tyr Val Ile Ile Asp Trp His
 130 135 140
 tca cat tct gca gaa aac aat cgc agt gaa gcc att cgc ttc ttc cag 1922
 Ser His Ser Ala Glu Asn Asn Arg Ser Glu Ala Ile Arg Phe Phe Gln
 145 150 155

gaa atg gcg cgc aaa tat ggc aac aag ccg aat gtc att tat gaa atc Glu Met Ala Arg Lys Tyr Gly Asn Lys Pro Asn Val Ile Tyr Glu Ile 160 165 170	1970
tac aac gag ccg ctt cag gtt tca tgg agc aat acc att aaa cct tat Tyr Asn Glu Pro Leu Gln Val Ser Trp Ser Asn Thr Ile Lys Pro Tyr 175 180 185	2018
gcc gaa gcc gtg att tcc gcc att cgc gcc att gac ccg gat aac ctg Ala Glu Ala Val Ile Ser Ala Ile Arg Ala Ile Asp Pro Asp Asn Leu 190 195 200 205	2066
att att gtc ggt acg ccc agt tgg tcg caa aac gtt gat gaa gcg tcg Ile Ile Val Gly Thr Pro Ser Trp Ser Gln Asn Val Asp Glu Ala Ser 210 215 220	2114
cgc gat cca atc aac gcc aag aat atc gcc tat acg ctg cat ttc tac Arg Asp Pro Ile Asn Ala Lys Asn Ile Ala Tyr Thr Leu His Phe Tyr 225 230 235	2162
gcg gga acc cat ggt gag tca tta cgc act aaa gcc cgc cag gcg tta Ala Gly Thr His Gly Glu Ser Leu Arg Thr Lys Ala Arg Gln Ala Leu 240 245 250	2210
aat aac ggt att gcg ctt ttc gtc acc gag tgg ggc gcc gtt aac gcg Asn Asn Gly Ile Ala Leu Phe Val Thr Glu Trp Gly Ala Val Asn Ala 255 260 265	2258
gac ggc aat ggc gga gtg aac cag aca gat acc gac gcc tgg gta acg Asp Gly Asn Gly Gly Val Asn Gln Thr Asp Thr Asp Ala Trp Val Thr 270 275 280 285	2306
ttc atg cgt gac aac aac atc agc aac gca aac tgg gcg tta aat gat Phe Met Arg Asp Asn Asn Ile Ser Asn Ala Asn Trp Ala Leu Asn Asp 290 295 300	2354
aaa agc gaa ggg gca tca acc tat tat ccg gac tct aaa aac ctg acc Lys Ser Glu Gly Ala Ser Thr Tyr Tyr Pro Asp Ser Lys Asn Leu Thr 305 310 315	2402
gag tcg ggt aaa ata gta aaa tcg atc att caa agc tgg cca tat aaa Glu Ser Gly Lys Ile Val Lys Ser Ile Ile Gln Ser Trp Pro Tyr Lys 320 325 330	2450
gcg ggc agc gcc gcc agt aca aca acc gat cag tca acc gat acc acc Ala Gly Ser Ala Ala Ser Thr Thr Thr Asp Gln Ser Thr Asp Thr Thr 335 340 345	2498
atg gca cca ccg ttg acg aac cga cca caa ccg aca cac cgg caa acc Met Ala Pro Pro Leu Thr Asn Arg Pro Gln Pro Thr His Arg Gln Thr 350 355 360 365	2546
gct gat tgc tgc aat gcc aac gtt tac ccc aac tgg gtt agc aaa gac Ala Asp Cys Cys Asn Ala Asn Val Tyr Pro Asn Trp Val Ser Lys Asp 370 375 380	2594

tgg gcg ggc cgg cag cga ctc ata acg aag cag gcc aat cga tcg tct	2642
Trp Ala Gly Arg Gln Arg Leu Ile Thr Lys Gln Ala Asn Arg Ser Ser	
385 390 395	
aca aag gga acc tgt ata ccg caa act ggt aca ctt cat ccg ttc cgg	2690
Thr Lys Gly Thr Cys Ile Pro Gln Thr Gly Thr Leu His Pro Phe Arg	
400 405 410	
gca gcg att cct cct ggg cac agg ttg gta gct gta act aat tga	2735
Ala Ala Ile Pro Pro Gly His Arg Leu Val Ala Val Thr Asn	
415 420 425	
ttaatctttt ccccccaaa ataacagggc tgcgattgca gcctgatacg caacattcca	2795
ttacttaatt gcgttcaaaa gcgccc aaat ccggtgcgct gccttgtaac taatatgatt	2855
tctctttcgt acccgcgta atcagctttg agttagccga cagacggaac agcgaggttg	2915
ccggcaacgt gccgtcatta tcacgagata cggtagccag cgagggtgtcc aggctgacga	2975
atcggaacgc gaagccgctg tccgtatcca tgagttgact cgcacccgca ttactgacgc	3035
ttgcagaagc agacagagac acgttggtgc ggaagtaatg tttctgtcct gactggacgt	3095
tgctcccgaa agcataatta atgccgtttt tatatgacgt gttatttatt accgtacgcc	3155
gccgcgttat tggttctggtc aaaacctttg ctacggttgc caaacgcgac gcaacgggta	3215
atgcgatgat tgccgaccgc tggttcctcc cagtttgaac ccgttggcat tgccggcgaa	3275
cgcgctnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3335
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3395
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3455
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3515
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3575
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3635
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3695
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3755
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3815
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3875
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3935
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	3995
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	4055
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	4115

nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4175
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4235
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnngatc	ctctagagtc	4295
gacctgcagg	aattcaactgg	ccgtcgtttt	acaacgtcgt	gactgggaaa	accttggcgt	4355
tacccaactt	aatcgccctg	cagcacatcc	ccctttcgcc	agctggcgta	atagcgaaga	4415
ggcccgacc	gatcgccctt	cccaacagtt	gcgcagcctg	aatggcgaat	ggcgccctgat	4475
gcggtatttt	ctccttacgc	atctgtgcgg	tatttcacac	cgcataggcg	cgcttatggt	4535
gcactctcag	tacaatctgc	tctgatgccg	catagttaag	ccagccccga	cacccgccaa	4595
cacccgctga	cgcgccctga	cgggcttgtc	tgctcccggc	atccgcttac	agacaagctg	4655
tgaccgtctc	ggggagctgc	atgtgtcaga	ggttttcacc	gtcatcaccg	aaacgcgcga	4715
gacgaaaggg	cctcgtgata	cgcctatttt	tatagggttaa	tgtcatgata	ataatggttt	4775
cttagacgtc	aggtggcact	tttcggggaa	atgtgcgcgg	aacccttatt	tgtttatttt	4835
tctaaaatac	ttcaaatacg	tatccgctca	tgagacaata	accttgataa	atgcttcaat	4895
aatataggaa	aaggaagagt	atg agt att	caa cat ttc	cgt gtc gcc	ctt att	4948
		Met Ser Ile	Gln His Phe	Arg Val Ala	Leu Ile	
		430		435		
ccc ttt ttt	gcg gca ttt	tgc ctt cct	ggt ttt gct	cac cca gaa	acg	4996
Pro Phe Phe	Ala Ala Phe	Cys Leu Pro	Val Phe Ala	His Pro Glu	Thr	
440		445		450		
ctg gtg aaa	gta aaa gat	gct gaa gat	cag ttg ggt	gca cga gtg	ggg	5044
Leu Val Lys	Val Lys Asp	Ala Glu Asp	Gln Leu Gly	Ala Arg Val	Gly	
455		460	465		470	
tac atc gaa	ctg gat ctc	aac agc ggt	aag atc ctt	gag agt ttt	cgc	5092
Tyr Ile Glu	Leu Asp Leu	Asn Ser Gly	Lys Ile Leu	Glu Ser Phe	Arg	
	475		480		485	
ccc gaa gaa	cgt ttt cca	atg atg agc	act ttt aaa	ggt ctg cta	tgt	5140
Pro Glu Glu	Arg Phe Pro	Met Met Ser	Thr Phe Lys	Val Leu Leu	Cys	
	490		495	500		
ggc gcg gta	tta tcc cgt	att gac gcc	ggg caa gag	caa ctc ggt	cgc	5188
Gly Ala Val	Leu Ser Arg	Ile Asp Ala	Gly Gln Glu	Gln Leu Gly	Arg	
	505	510		515		
cgc ata cac	tat tct cag	aat gac ttg	ggt gag tac	tca cca gtc	aca	5236
Arg Ile His	Tyr Ser Gln	Asn Asp Leu	Val Glu Tyr	Ser Pro Val	Thr	
	520	525	530			
gaa aag cat	ctt acg gat	ggc atg aca	gta aga gaa	tta tgc agt	gct	5284
Glu Lys His	Leu Thr Asp	Gly Met Thr	Val Arg Glu	Leu Cys Ser	Ala	
535		540	545		550	

gcc ata acc atg agt gat aac act gcg gcc aac tta ctt ctg aca acg Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr 555 560 565	5332
atc gga gga ccg aag gag cta acc gct ttt ttg cac aac atg ggg gat Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp 570 575 580	5380
cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa gcc ata His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile 585 590 595	5428
cca aac gac gag cgt gac acc acg atg cct gta gca atg gca aca acg Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr 600 605 610	5476
ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg caa Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln 615 620 625 630	5524
caa tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu 635 640 645	5572
cgc tcg gcc ctt ccg gct ggc tgg ttt att gct gat aaa tct gga gcc Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala 650 655 660	5620
ggg gag cgt ggg tct cgc ggt atc att gca gca ctg ggg cca gat ggt Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly 665 670 675	5668
aag ccc tcc cgt atc gta gtt atc tac acg acg ggg agt cag gca act Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr 680 685 690	5716
atg gat gaa cga aat aga cag atc gct gag ata ggt gcc tca ctg att Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile 695 700 705 710	5764
aag cat tgg taa ctgtcagacc aagtttactc atatatactt tagattgatt Lys His Trp	5816
taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga	5876
ccaaaatccc ttaacgtgag ttttcggtcc actgagcgtc agaccccgta gaaaagatca	5936
aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac	5996
caccgctacc agcgggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg	6056
taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag	6116
gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac	6176
cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt	6236

taccggataa ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg	6296
agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc	6356
ttcccgaagg gagaaaggcg gacagggtatc cggttaagcgg cagggtcggg acaggagagc	6416
gcacgagggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc	6476
acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa	6536
acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt	6596
tctttcctgc gttatccctt gattctgtgg ataaccgtat taccgccttt gaggagctg	6656
ataccgctcg ccgcagccga acgaccgagc gcagcagtc agtgagcgag gaagcggcgc	6716
gccagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta	6776
atgcagctgg cagcacaggt ttcccgaactg gaaagcgggc agtgagcgca acgcaattaa	6836
tgtgagttag ctactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat	6896
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta	6956
cgccaagctt gcatgccaat tctcatgttt gacagcttat catcgataag ctttaatgcg	7016
gtagtttatc acagttaaat tgctaacgca gtcaggcacc gtgt atg aaa tct aac	7072
	Met Lys Ser Asn
	715
aat gcg ctc atc gtc atc ctc ggc acc gtc acc ctg gat gct gta ggc	7120
Asn Ala Leu Ile Val Ile Leu Gly Thr Val Thr Leu Asp Ala Val Gly	
720 725 730	
ata ggc ttg gtt atg ccg gta ctg ccg ggc ctc ttg cgg gat atc gtc	7168
Ile Gly Leu Val Met Pro Val Leu Pro Gly Leu Leu Arg Asp Ile Val	
735 740 745	
cat tcc gac agc atc gcc agt cac tat ggc gtg ctg cta gcg cta tat	7216
His Ser Asp Ser Ile Ala Ser His Tyr Gly Val Leu Leu Ala Leu Tyr	
750 755 760 765	
gcg ttg atg caa ttt cta tgc gca ccc gtt ctc gga gca ctg tcc gac	7264
Ala Leu Met Gln Phe Leu Cys Ala Pro Val Leu Gly Ala Leu Ser Asp	
770 775 780	
cgc ttt ggc cgc cgc cca gtc ctg ctc gct tcg cta ctt gga gcc act	7312
Arg Phe Gly Arg Arg Pro Val Leu Leu Ala Ser Leu Leu Gly Ala Thr	
785 790 795	
atc gac tac gcg atc atg gcg acc aca ccc gtc ctg tgg atc ctc tac	7360
Ile Asp Tyr Ala Ile Met Ala Thr Thr Pro Val Leu Trp Ile Leu Tyr	
800 805 810	
gcc gga cgc atc gtg gcc ggc atc acc ggc gcc aca ggt gcg gtt gct	7408
Ala Gly Arg Ile Val Ala Gly Ile Thr Gly Ala Thr Gly Ala Val Ala	
815 820 825	

ggc gcc tat atc gcc gac atc acc gat ggg gaa gat cgg gct cgc cac Gly Ala Tyr Ile Ala Asp Ile Thr Asp Gly Glu Asp Arg Ala Arg His 830 835 840 845	7456
ttc ggg ctc atg agc gct tgt ttc ggc gtg ggt atg gtg gca ggc ccc Phe Gly Leu Met Ser Ala Cys Phe Gly Val Gly Met Val Ala Gly Pro 850 855 860	7504
gtg gcc ggg gga ctg ttg ggc gcc atc tcc ttg cat gca cca ttc ctt Val Ala Gly Gly Leu Leu Gly Ala Ile Ser Leu His Ala Pro Phe Leu 865 870 875	7552
gcg gcg gcg gtg ctc aac ggc ctc aac cta cta ctg ggc tgc ttc cta Ala Ala Ala Val Leu Asn Gly Leu Asn Leu Leu Leu Gly Cys Phe Leu 880 885 890	7600
atg cag gag tcg cat aag gga gag cgt cga ccg atg ccc ttg aga gcc Met Gln Glu Ser His Lys Gly Glu Arg Arg Pro Met Pro Leu Arg Ala 895 900 905	7648
ttc aac cca gtc agc tcc ttc cgg tgg gcg cgg ggc atg act atc gtc Phe Asn Pro Val Ser Ser Phe Arg Trp Ala Arg Gly Met Thr Ile Val 910 915 920 925	7696
gcc gca ctt atg act gtc ttc ttt atc atg caa ctc gta gga cag gtg Ala Ala Leu Met Thr Val Phe Phe Ile Met Gln Leu Val Gly Gln Val 930 935 940	7744
ccg gca gcg ctc tgg gtc att ttc ggc gag gac cgc ttt cgc tgg agc Pro Ala Ala Leu Trp Val Ile Phe Gly Glu Asp Arg Phe Arg Trp Ser 945 950 955	7792
gcg acg atg atc ggc ctg tcg ctt gcg gta ttc gga atc ttg cac gcc Ala Thr Met Ile Gly Leu Ser Leu Ala Val Phe Gly Ile Leu His Ala 960 965 970	7840
ctc gct caa gcc ttc gtc act ggt ccc gcc acc aaa cgt ttc ggc gag Leu Ala Gln Ala Phe Val Thr Gly Pro Ala Thr Lys Arg Phe Gly Glu 975 980 985	7888
aag cag gcc att atc gcc ggc atg gcg gcc gac gcg ctg ggc tac gtc Lys Gln Ala Ile Ile Ala Gly Met Ala Ala Asp Ala Leu Gly Tyr Val 990 995 1000 1005	7936
ttg ctg gcg ttc gcg acg cga ggc tgg atg gcc ttc ccc att atg Leu Leu Ala Phe Ala Thr Arg Gly Trp Met Ala Phe Pro Ile Met 1010 1015 1020	7981
att ctt ctc gct tcc ggc ggc atc ggg atg ccc gcg ttg cag gcc Ile Leu Leu Ala Ser Gly Gly Ile Gly Met Pro Ala Leu Gln Ala 1025 1030 1035	8026
atg ctg tcc agg cag gta gat gac gac cat cag gga cag ctt caa Met Leu Ser Arg Gln Val Asp Asp Asp His Gln Gly Gln Leu Gln 1040 1045 1050	8071

gga tcg ctc gcg gct	ctt acc agc cta act	tcg atc act gga ccg	8116
Gly Ser Leu Ala Ala	Leu Thr Ser Leu Thr	Ser Ile Thr Gly Pro	
1055	1060	1065	
ctg atc gtc acg gcg	att tat gcc gcc tcg	gcg agc aca tgg aac	8161
Leu Ile Val Thr Ala	Ile Tyr Ala Ala Ser	Ala Ser Thr Trp Asn	
1070	1075	1080	
ggg ttg gca tgg att	gta ggc gcc gcc cta	tac ctt gtc tgc ctc	8206
Gly Leu Ala Trp Ile	Val Gly Ala Ala Leu	Tyr Leu Val Cys Leu	
1085	1090	1095	
ccc gcg ttg cgt cgc	ggt gca tgg agc cgg	gcc acc tcg acc tga	8251
Pro Ala Leu Arg Arg	Gly Ala Trp Ser Arg	Ala Thr Ser Thr	
1100	1105		
atggaagccg gcggcacctc	gctaacggat tcaccactcc	aagaattgga gccaatcaat	8311
tcttgccggag aactgtgaat	gcgcaaacca acccttggca	gaacatatcc atcgcgtccg	8371
ccatctccag cagccgcacg	cggcgcatct cggggctcgac	tctagaggat ccccgcaacg	8431
ctgtcagcgc tttccagtta	aacggctcca acgtcgccat	aggtaattcc tcgcccggcc	8491
atacgatcgg gcaggtgccg	ttggctatcg ccgctgcctg	actcatcaca ctatcttccg	8551
ctgcatcgcg aagggttttg	accacttctt ccattctctc	gtgcgccgga tgccatgctc	8611
acgtacgcgg cttatcagat	agtcgggcag gccgtcggtc	cagcccaatg aggggaagct	8671
ggcgtggagc gatgccagca	cctgctcctc aacaccgtaa	tggccggcgg cgaacaggca	8731
ttcggcggtta agcgtttcca	gccctttaat catcacgtg	cggcacatct tgatagccga	8791
cacgctgcca acgtggttac	caccatagcg ggcgttacat	ccaagcgtgg tgagtaattc	8851
agcaattgcc tctgcctgtg	gtccccccgt caacagcggc	gttcggagtg cccctggggg	8911
gaccggcgcc atcaccgcta	catcgacata agcgcggggc	ttaaagcatt tggcagcctg	8971
acgcttggtc tgcggggcga	cggagttaag gtcaagaaaa	tactgcgtgt cggtcatcag	9031
cgggtgcagct tgtgaggcga	catccagggc ggatcccgcg	gtgacggtgg aaaatatgag	9091
ttcggcacct gtcaacgcgt	cagccaggga gattgccgcc	cgcacgcctc cccgatgcgc	9151
cttcgttatc atcgcatcgc	gctcaggacc ttgcagcttg	caatcccaga cggtgactgg	9211
gttcactttt gccagtgcac	ccgcaagaat gccacctgct	tcaccataac ctataaacgt	9271
tattgtcgtc ataacagctc	cttacgcggc cacacgtcgg	ccggaatgca aacgtcgccc	9331
gcgaacagaa gtcgcgccgt	acgcagcaga ccgcagcctg	ccaactgccc attatcatca	9391
agccggagcg ccacgctgaa	ttgggtaccg agctccgaat	tgggtaccga gctcgaatta	9451
attcgagctc ggtacccggg	gacnnnnnnn nnnnnnnnnn	nnnnnnnnnn nnnnnnnnnn	9511

9571	9571	9571	9571	9571	9571	9571
9631	9631	9631	9631	9631	9631	9631
9691	9691	9691	9691	9691	9691	9691
9751	9751	9751	9751	9751	9751	9751
9811	9811	9811	9811	9811	9811	9811
9871	9871	9871	9871	9871	9871	9871
9931	9931	9931	9931	9931	9931	9931
9991	9991	9991	9991	9991	9991	9991
10051	10051	10051	10051	10051	10051	10051
10111	10111	10111	10111	10111	10111	10111
10171	10171	10171	10171	10171	10171	10171
10231	10231	10231	10231	10231	10231	10231
10291	10291	10291	10291	10291	10291	10291
10351	10351	10351	10351	10351	10351	10351
10411	10411	10411	10411	10411	10411	10411
10471	10471	10471	10471	10471	10471	10471
10531	10531	10531	10531	10531	10531	10531
10591	10591	10591	10591	10591	10591	10591
10651	10651	10651	10651	10651	10651	10651
10711	10711	10711	10711	10711	10711	10711
10771	10771	10771	10771	10771	10771	10771
10831	10831	10831	10831	10831	10831	10831
10891	10891	10891	10891	10891	10891	10891
10951	10951	10951	10951	10951	10951	10951
11011	11011	11011	11011	11011	11011	11011
11071	11071	11071	11071	11071	11071	11071
11131	11131	11131	11131	11131	11131	11131
11191	11191	11191	11191	11191	11191	11191
11251	11251	11251	11251	11251	11251	11251

nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11311
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11371
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11431
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11491
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnn 11544

<210> 13

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 13

ctgttcggtt accaacac

18

<210> 14

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 14

gtgaatggga tcacgagt

18

<210> 15

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 15

accatcagca tcaacgccca acaacg

26

<210> 16

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 16

gactggatgg ttatccgaat aagagagagg

30

<210> 17
 <211> 11772
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic nucleotide construct

<220>
 <221> modified_base
 <222> (144)..(1143)
 <223> a, c, g, t, unknown, or other

<220>
 <221> modified_base
 <222> (5753)..(7567)
 <223> a, c, g, t, unknown, or other

<400> 17
 atcgatgaat tgatccgaag ttcctattct ctagaaagta taggaacttc gaattgtcga 60
 caagcttgat ctggcttata gaaattaata cgactcacta tagggagacc ggaattcccc 120
 tgcaggtcga ctctagagga tcannnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 180
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 240
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 300
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 360
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 420
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 480
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 540
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 600
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 660
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 720
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 780
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 840
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 900
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 960
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1020
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1080

nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	1140
nnnagcgcgt tcgccggcaa tgccaacggg ttcaaactgg gaggaaccag cggtcggcaa	1200
tcatcgcatt acccgttgcg tcgcgtttgg caacgtgagc aaaggttttg accagaacaa	1260
taacgcggcg gcgtacggta ataaataaca cgtcatataa aaacggcatt aattatgctt	1320
tcgggagcaa cgtccagtca ggacagaaac attacttccg caacaacgtg tctctgtctg	1380
cttctgcaac ggtcagtaat gcggatgcga gtcaactcat ggatacggac agcggcttcc	1440
gcgtccgatt cgtcagcctg gacacctcgc tggctaccgt atctcgtgat aatgacggca	1500
cgttgccggc aacctcgctg ttccgtctgt cggctaactc aaagctgatt aacgcgggta	1560
cgaaagagaa atcatattag ttacaaggca gcgcaccgga tttgggcgct tttgaacgca	1620
attaagtaat ggaatgttgc gtatcaggct gcaatcgcag ccctgttatt ttgggggtga	1680
aaagattaat caattagtta cagctaccaa cctgtgcca ggaggaatcg ctgcccgga	1740
cggatgaagt gtaccagttt gcggtataca ggttcccttt gtagacgac gattggcctg	1800
cttcgttatg agtcgctgcc ggcccgcca gtctttgcta acccagttgg ggtaaacgtt	1860
ggcattgcag caatcagcgg tttgccggtg tgtcggttgt ggtcggttcg tcaacgggtg	1920
tgccatggtg gtatcggttg actgatcggg tgttgactcg gcggcgctgc ccgctttata	1980
tggccagctt tgaatgatcg attttactat ttacccgac tcggtcagggt ttttagagtc	2040
cggataatag gttgatgccc cttcgctttt atcatttaac gccagtttg cgttgctgat	2100
gttggtgtca cgcataacg ttaccaggc gtcggtatct gtctggttca ctccgccatt	2160
gccgtccgcg ttaacggcgc cccactcggg gacgaaaagc gcaataccgt tatttaacgc	2220
ctggcgggct ttagtgcgta atgactcacc atgggttccc gcgtagaaat gcagcgtata	2280
ggcgatatto ttggcgttga ttggatcgcg cgacgcttca tcaacgtttt gcgaccaact	2340
gggcgtaccg acaataatca ggttatccgg gtcaatggcg cgaatggcg aaatcacggc	2400
ttcggcataa ggtttaatgg tattgtcca tgaaacctga agcggctcgt tgtagatttc	2460
ataaatgaca ttcggcttgt tgccatattt gcgcgccatt tcctggaaga agcgaatggc	2520
ttcactgcga ttgttttctg cagaatgtga gtgccagtca ataacacat acatatcgtt	2580
ggcgattgcg gcatccacca ctctttcaac tttggccttg ttgccagccg ggtcctgcag	2640
ataaccaccg ctttcctgaa cgcccatagc ggcgcgaaca atgctggatt tccagtcttt	2700
tttcagcgac gcaacggtat cggctgtgta gaatttttcc ccacccaac cattattact	2760
ccagaataag ctgttgccgg caaaactttt ggctttttca cctgcgtaga ttttattgcc	2820

gctaacggat aacgggtcaa cactcgccca ggcattactg gaaaggcagg caaggcttaa	2880
tcctaaacag gcacaagata gaaacagttt ttacgtggg gcgtgttttt ggctatcgat	2940
gactggatgg ttatccgaat aagagagagg cataaatgaa tctccatttc atagactcta	3000
gaggatcacg aaaatcaggc agcgcttaaa tctgcaaaa caccggaggc tgcttcatta	3060
tggcagcgca aggcaatttc agtataaagg gcatcacgat cgaaatcaat gctgccgccc	3120
tgaatggctt ccatcaggtc ttctgcgccg acaacttctg caccagcagc tttagcttct	3180
tcagctttcg ggccacgggc gaaaacaccg acgcgaacat ttttaccggg gcccttcggc	3240
aaggtaacaa caccacgaac catctgatca aggtattaag gccgccggac aggtaaaatg	3300
gttatggcta accgatgccg atattgcgca tcaggatgat accttgaaaa tgttggttgc	3360
tcatgccgaa aaagacggat tgatttttca ttccctgatg gcgcgactaa gttgtaacaa	3420
ttttgcagaa tgggcgctta ttccggcctt cgtctttttc ttccagatgt tgtatccttt	3480
tcggcatatc aataatccca agcataaaac ggccgggtgct gccgggtggat gcatgttagc	3540
cgatcatcat ctgtgctggg ttggataaac cgttactaat ctggcgcgcg acccgataga	3600
cagcggcttc cccagcccc agtgagctac caatagcaag ggtgccaccc tgctggaaaa	3660
tggagtctag gatttcatta aaactgacgc ttaaagtga tttccacatc ccttcaacgg	3720
gcgtttttct ctgaaatatt tttcttatcg gaaatgctcg ctccgtatat tggtgaaaga	3780
gataaatgcc agcgtaacta caggtgacaa aaagcgtgaa ttgcgtcagg caccatataa	3840
acaaaaaata ccccaaaggc atatgaagcc aataaccaat gccacaacct ccggttctga	3900
cgcaggtcgt aataaactca taaatagtga ccagtttaaa gcggttacac agccgcaaca	3960
tcccggttgg accagccgat attcataaaa agtataatca gcatacaaag caaggcatct	4020
ggcttgacct cggcaggcca tccaatctt gaagtgccta aaatcaagat accgcctaac	4080
cctaccaaca tacctatcgc cgcactaaaa aaatcggctc tgatgcaaaa ggcaaggaca	4140
tcatcaaatt ggtaaaaatc tttttcctga aaagcttttg aaccgtaatg cagcaagggt	4200
tgccatgact gaaaacgggt aatatccgaa atcagggttag caaaagtcgt aatcaaaagc	4260
atgacaccaa acaggttcaa tccaatgtc tgggatgtcc atgcgacata gacaaaactg	4320
cataccgcat tcaaaacacg gcctgtaatc agaatgccgg tattgccaat gatacgggaa	4380
aagacagatc gaatttgatg ccgtggataa attgccgggt gatcgatccc cgggtaccga	4440
gctcgaattc cgagcttggc gcgcctatgc ggtgtgaaat accgcacaga tgcgtaagga	4500
gaaaataccg catcaggcgc cattcgccat tcaggctgcg caactgttgg gaagggcgat	4560

nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6360
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6420
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6480
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6540
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6600
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6660
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6720
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6780
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6840
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6900
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	6960
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7020
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7080
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7140
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7200
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7260
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7320
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7380
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7440
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7500
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	7560
nnnnnnnnngaa	ttcgagctcg	gtaccgggg	atcgacataa	ccgataggct	ctgcattgat	7620
ggactgaaag	gtttcgacgg	cggttgtgtg	ggttttgctt	gcccatctgg	cggcatgaat	7680
agtgtcattc	atgacgatcc	agttcgatat	tcaacagacc	gtctttgtaa	tcggcaccga	7740
caattttgat	caataaagcg	tttgacctga	tgcattgagg	taaatccatt	cgttcggttg	7800
ttcttttctg	attacctgtc	ctgttaacct	gtggatatag	aaggctcggtt	caatgagtag	7860
tattctgacg	catctgacaa	ttggttccaa	tgacctgaag	aaggcgcgca	tcttttatga	7920
tgctgttttg	gaaccgttgg	gtatcaaact	tattcgcgag	gtcgaaggac	agcgttttgc	7980
ctatggtaaa	gacggcgaag	aaggacgcat	catcattgta	aagcctatta	atgggtgaagc	8040

cgctaccgct ggaaatggta tcactatcgg tttggcagcg ccttctgatg aagctgtcga	8100
tgctttttat aaagcaggct tggctaattg cggttaaggat gccggagAAC cggggcctcg	8160
tccggctgct aataattctc ggggtgccta tttatatgac cctgaaggca ataaaatctg	8220
cgctttcaat tttaaataag atttcttttg tgcagggtta ttcaaaatag ccctgcattt	8280
tcagtattat agcggccatt atggcttttg ccttgataaa aaatttatca gggctgtttt	8340
tcgtgatgaa tatttttgat ttttcaagaa aagcctgata tcttccaaca tctttctttg	8400
tatataaatg gagcgagcta tggcgcgct aactgtcgaa gactgtatcg ataaagttca	8460
taatcgtttc gatttgatcc tctagagtca acctgcttgt tactcgtgat ccattcaca	8520
aggcggaatt aattcgccct tctgttccgt taccaacact tgagccggag gcataatggg	8580
aaagccaatg tggcgttgtt gggcgttgat gctgatggtg tggttcagtg cgtcggctac	8640
ggcggcgaac ggctgggaaa tctataaaag cgttttcatg accacggacg ggcgcattca	8700
ggataccggc aataagaatg tcagccacac cgaaggtcag ggattcgcca tgctgatggc	8760
ggtgcattac gatgaccgca tcgcgttcga taacctgtgg aactggacgc aaagccacct	8820
gcggaacacg accagcggct tgttctactg gcgttacgat ccgtcggcgg ccaatccggt	8880
ggtggataag aacaacgcct cggatggcga tgtgctgatt gcctgggcgt tgttaaaagc	8940
gggaaataag tggcaggaca accgttacct gcaggcgtcg gacagcatcc agaaagcgat	9000
catcgccagc aatatcattc agtttgccgg ccgcaccgtg atgttgcccg gcgcctatgg	9060
tttcaacaag aacagctatg tgatccttaa cccgtcgtat ttctgttcc cggcctggcg	9120
cgactttgct aaccgcagcc atcttcaggt gtggcggcaa ctgattgacg acagcctgtc	9180
attggtcggg gaaatgcgtt tcggtcaggt cgggctgccg acggactggg cggcgctgaa	9240
cgcggatggc tcgatggcgc cggcgacggc ctggccgtcg cgtttcagtt acgacgcat	9300
tcgtatcccg ctgtatttgt actggtatga cgccaaaacc acggcgctgg tgccgttcca	9360
gctgtactgg cgtaactatc cccgcctgac gacgccggcc tgggttgatg tgctgagcag	9420
taacaccgcg acttacaata tgcagggcgg tttgctggcg gtgcgcgacc tgacgatggg	9480
caacctcgac gggctcagcg atctgccagg cgcacggaa gattactact cgtcgagcct	9540
gcgcctgctg gtgatgttgg cgcgcggtaa ataaccttat tcttgcggtg cacatggcga	9600
ggacgatgtc cttgccattt tccccacttt tatccctctg aatggcgtgt ttttcacgct	9660
ttgttaacct gcttgttact cgtgatccca ttcacaaggg cgaattgacc tgcaggcatg	9720
caagcttggc gtaatcatgg tcatagctgt ttctgtgtg aaattgttat ccgctcaca	9780

ttccacacaa	catacgagcc	ggaagcataa	agtgtaaagc	ctgggggtgcc	taatgagtga	9840
gctaactcac	attaattgcg	ttgcgctcac	tgcccgcctt	ccagtcggga	aacctgtcgt	9900
gccagctgca	ttaatgaatc	ggccaacgcg	cggggagagg	cggtttgcgt	attgggcgct	9960
cttccgctgg	cgcgccaggt	cgactctaga	ggatccccgg	ggaagatctt	ccggaagatc	10020
ttcccagagc	cgaattaatt	ccgcgatgaa	ttgatccccg	aagttcctat	tctctagaaa	10080
gtataggaac	tcgaattggt	cgacaagcta	gcttgcacgc	aagcttgat	tctatagtgt	10140
cacctaaatc	gtatgtgtat	gatacataag	gttatgtatt	aattgtagcc	gcgttctaac	10200
gacaatatgt	acaagcctaa	ttgtgtagca	tctggcttac	tgaagcagac	cctatcatct	10260
ctctcgtaaa	ctgccgtcag	agtcggtttg	gttggaacga	ccttctgagt	ttctggtaac	10320
gccgttccgc	accccgga	tggtcagcga	accaatcagc	agggtcatcg	ctagcccatg	10380
gctaattctg	tcagccgtta	agtgttcctg	tgctactgaa	aattgctttg	agaggctcta	10440
agggttctc	agtgcgttac	atccctggct	tgttgtccac	aaccgttaaa	ccttaaaagc	10500
tttaaaagcc	ttatatattc	ttttttttct	tataaaactt	aaaaccttag	aggctattta	10560
agttgctgat	ttatattaat	tttattgttc	aaacatgaga	gcttagtacg	tgaaacatga	10620
gagcttagta	cgtagccat	gagagcttag	tacgttagcc	atgagggttt	agttcgtaaa	10680
acatgagagc	ttagtacgtt	aaacatgaga	gcttagtacg	tgaaacatga	gagcttagta	10740
cgtactatca	acaggttgaa	ctgcggatct	tgcgcccgca	aaaattaaaa	atgaagtttt	10800
gacggtatcg	aacccagag	tcccgtcag	aagaactcgt	caagaaggcg	atagaaggcg	10860
atgcgctgcg	aatcgggagc	ggcgataccg	taaagcacga	ggaagcggtc	agccattcgc	10920
ccgccaagct	cttcagcaat	atcacgggta	gccaacgcta	tgctctgata	gcggtccgcc	10980
acacccagcc	ggccacagtc	gatgaatcca	gaaaagcggc	cattttccac	catgatattc	11040
ggcaagcagg	catcgccatg	ggtcacgacg	agatcctcgc	cgtcgggcat	ccgcgccttg	11100
agcctggcga	acagttcggc	tggcgcgagc	ccctgatgct	cttcgtccag	atcatcctga	11160
tcgacaagac	cggcttccat	ccgagtacgt	gctcgcctga	tgcatgtttt	cgtttggtgg	11220
tcgaatgggc	aggtagccgg	atcaagcgta	tgacgccgcc	gcattgcac	agccatgatg	11280
gatactttct	cggcaggagc	aaggtgagat	gacaggagat	cctgccccgg	cacttcgccc	11340
aatagcagcc	agtccttcc	cgcttcagt	acaacgtcga	gcacagctgc	gcaagggaacg	11400
cccgtcgtgg	ccagccacga	tagccgcgct	gcctcgtctt	ggagttcatt	cagggcaccg	11460
gacaggtcgg	tcttgacaaa	aagaaccggg	cgccctgcg	ctgacagccg	gaacacggcg	11520

gcatcagagc agccgattgt ctgttggtgcc cagtcatagc cgaatagcct ctccacccaa 11580
 gcggccggag aacctgcgtg caatccatct tgttcaatca tgcgaaacga tcctcatcct 11640
 gtctcttgat ccactagatt attgaagcat ttatcagggt tattgtctca tgagcggata 11700
 catatttgaa tgtattttaga aaaataaaca aatagggggt ccgcgcacat ttccccgaaa 11760
 agtgccacct gc 11772

<210> 18
 <211> 45
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> 18
 atatttttga tttttcaaga aaagcctgat atcttccaac atctt 45

<210> 19
 <211> 45
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> 19
 gatttgatcc tctagagtca acctgcttgt tactcgtgat cccat 45

<210> 20
 <211> 44
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> 20
 gagtcaacct gcttggttact cgtgatccca ttcacaaggg cgaa 44

<210> 21
 <211> 42
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 21

ttactcgtga tcccatcac aaggcgcaat taattcgccc tt

42

<210> 22

<211> 427

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
amino acid sequence

<400> 22

Met Pro Leu Ser Tyr Ser Asp Asn His Pro Val Ile Asp Ser Gln Lys
1 5 10 15His Ala Pro Arg Lys Lys Leu Phe Leu Ser Cys Ala Cys Leu Gly Leu
20 25 30Ser Leu Ala Cys Leu Ser Ser Asn Ala Trp Ala Ser Val Glu Pro Leu
35 40 45Ser Val Ser Gly Asn Lys Ile Tyr Ala Gly Glu Lys Ala Lys Ser Phe
50 55 60Ala Gly Asn Ser Leu Phe Trp Ser Asn Asn Gly Trp Gly Gly Glu Lys
65 70 75 80Phe Tyr Thr Ala Asp Thr Val Ala Ser Leu Lys Lys Asp Trp Lys Ser
85 90 95Ser Ile Val Arg Ala Ala Met Gly Val Gln Glu Ser Gly Gly Tyr Leu
100 105 110Gln Asp Pro Ala Gly Asn Lys Ala Lys Val Glu Arg Val Val Asp Ala
115 120 125Ala Ile Ala Asn Asp Met Tyr Val Ile Ile Asp Trp His Ser His Ser
130 135 140Ala Glu Asn Asn Arg Ser Glu Ala Ile Arg Phe Phe Gln Glu Met Ala
145 150 155 160

Arg Lys Tyr Gly Asn Lys Pro Asn Val Ile Tyr Glu Ile Tyr Asn Glu
 165 170 175

Pro Leu Gln Val Ser Trp Ser Asn Thr Ile Lys Pro Tyr Ala Glu Ala
 180 185 190

Val Ile Ser Ala Ile Arg Ala Ile Asp Pro Asp Asn Leu Ile Ile Val
 195 200 205

Gly Thr Pro Ser Trp Ser Gln Asn Val Asp Glu Ala Ser Arg Asp Pro
 210 215 220

Ile Asn Ala Lys Asn Ile Ala Tyr Thr Leu His Phe Tyr Ala Gly Thr
 225 230 235 240

His Gly Glu Ser Leu Arg Thr Lys Ala Arg Gln Ala Leu Asn Asn Gly
 245 250 255

Ile Ala Leu Phe Val Thr Glu Trp Gly Ala Val Asn Ala Asp Gly Asn
 260 265 270

Gly Gly Val Asn Gln Thr Asp Thr Asp Ala Trp Val Thr Phe Met Arg
 275 280 285

Asp Asn Asn Ile Ser Asn Ala Asn Trp Ala Leu Asn Asp Lys Ser Glu
 290 295 300

Gly Ala Ser Thr Tyr Tyr Pro Asp Ser Lys Asn Leu Thr Glu Ser Gly
 305 310 315 320

Lys Ile Val Lys Ser Ile Ile Gln Ser Trp Pro Tyr Lys Ala Gly Ser
 325 330 335

Ala Ala Ser Thr Thr Thr Asp Gln Ser Thr Asp Thr Thr Met Ala Pro
 340 345 350

Pro Leu Thr Asn Arg Pro Gln Pro Thr His Arg Gln Thr Ala Asp Cys
 355 360 365

Cys Asn Ala Asn Val Tyr Pro Asn Trp Val Ser Lys Asp Trp Ala Gly
 370 375 380

Arg Gln Arg Leu Ile Thr Lys Gln Ala Asn Arg Ser Ser Thr Lys Gly
 385 390 395 400

Thr Cys Ile Pro Gln Thr Gly Thr Leu His Pro Phe Arg Ala Ala Ile
 405 410 415

Pro Pro Gly His Arg Leu Val Ala Val Thr Asn
 420 425

<210> 23

<211> 286

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
 amino acid sequence

<400> 23

Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
 1 5 10 15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
 20 25 30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
 35 40 45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
 50 55 60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
 65 70 75 80

Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
 85 90 95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
 100 105 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
 115 120 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
 130 135 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
 145 150 155 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
 165 170 175

Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
 180 185 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
 195 200 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
 210 215 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
 225 230 235 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
 245 250 255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
 260 265 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
 275 280 285

<210> 24

<211> 396

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
 amino acid sequence

<400> 24

Met Lys Ser Asn Asn Ala Leu Ile Val Ile Leu Gly Thr Val Thr Leu
 1 5 10 15

Asp Ala Val Gly Ile Gly Leu Val Met Pro Val Leu Pro Gly Leu Leu
 20 25 30

Arg Asp Ile Val His Ser Asp Ser Ile Ala Ser His Tyr Gly Val Leu
 35 40 45

Leu Ala Leu Tyr Ala Leu Met Gln Phe Leu Cys Ala Pro Val Leu Gly
 50 55 60

Ala Leu Ser Asp Arg Phe Gly Arg Arg Pro Val Leu Leu Ala Ser Leu
 65 70 75 80

Leu Gly Ala Thr Ile Asp Tyr Ala Ile Met Ala Thr Thr Pro Val Leu
 85 90 95

Trp Ile Leu Tyr Ala Gly Arg Ile Val Ala Gly Ile Thr Gly Ala Thr
 100 105 110

Gly Ala Val Ala Gly Ala Tyr Ile Ala Asp Ile Thr Asp Gly Glu Asp
 115 120 125

Arg Ala Arg His Phe Gly Leu Met Ser Ala Cys Phe Gly Val Gly Met
 130 135 140

Val Ala Gly Pro Val Ala Gly Gly Leu Leu Gly Ala Ile Ser Leu His
 145 150 155 160

Ala Pro Phe Leu Ala Ala Ala Val Leu Asn Gly Leu Asn Leu Leu Leu
 165 170 175

Gly Cys Phe Leu Met Gln Glu Ser His Lys Gly Glu Arg Arg Pro Met
 180 185 190

Pro Leu Arg Ala Phe Asn Pro Val Ser Ser Phe Arg Trp Ala Arg Gly
 195 200 205

Met Thr Ile Val Ala Ala Leu Met Thr Val Phe Phe Ile Met Gln Leu
 210 215 220

Val Gly Gln Val Pro Ala Ala Leu Trp Val Ile Phe Gly Glu Asp Arg
 225 230 235 240

Phe Arg Trp Ser Ala Thr Met Ile Gly Leu Ser Leu Ala Val Phe Gly
 245 250 255

Ile Leu His Ala Leu Ala Gln Ala Phe Val Thr Gly Pro Ala Thr Lys
 260 265 270

Arg Phe Gly Glu Lys Gln Ala Ile Ile Ala Gly Met Ala Ala Asp Ala
 275 280 285

Leu Gly Tyr Val Leu Leu Ala Phe Ala Thr Arg Gly Trp Met Ala Phe
 290 295 300

Pro Ile Met Ile Leu Leu Ala Ser Gly Gly Ile Gly Met Pro Ala Leu
 305 310 315 320

Gln Ala Met Leu Ser Arg Gln Val Asp Asp Asp His Gln Gly Gln Leu
 325 330 335

Gln Gly Ser Leu Ala Ala Leu Thr Ser Leu Thr Ser Ile Thr Gly Pro
 340 345 350

Leu Ile Val Thr Ala Ile Tyr Ala Ala Ser Ala Ser Thr Trp Asn Gly
 355 360 365

Leu Ala Trp Ile Val Gly Ala Ala Leu Tyr Leu Val Cys Leu Pro Ala
 370 375 380

Leu Arg Arg Gly Ala Trp Ser Arg Ala Thr Ser Thr
 385 390 395